Learning Swing-free Trajectories for UAVs with a Suspended Load

Aleksandra Faust[1], Ivana Palunko[2],
Patricio Cruz[3], Rafael Fierro[3], and Lydia Tapia[1]

[1]Department of Computer Science,
University of New Mexico, USA

[2]Faculty of Electrical Engineering and Computing,
University of Zagreb, Croatia

[3]Department of Electrical and Computer Engineering, University of New Mexico, USA

THE UNIVERSITY of NEW MEXICO

ADAPTIVE MOTION PLANNING RESEARCH GROUP

ACROSS
Centre of Research Excellence for Advanced Cooperative Systems

Motivation

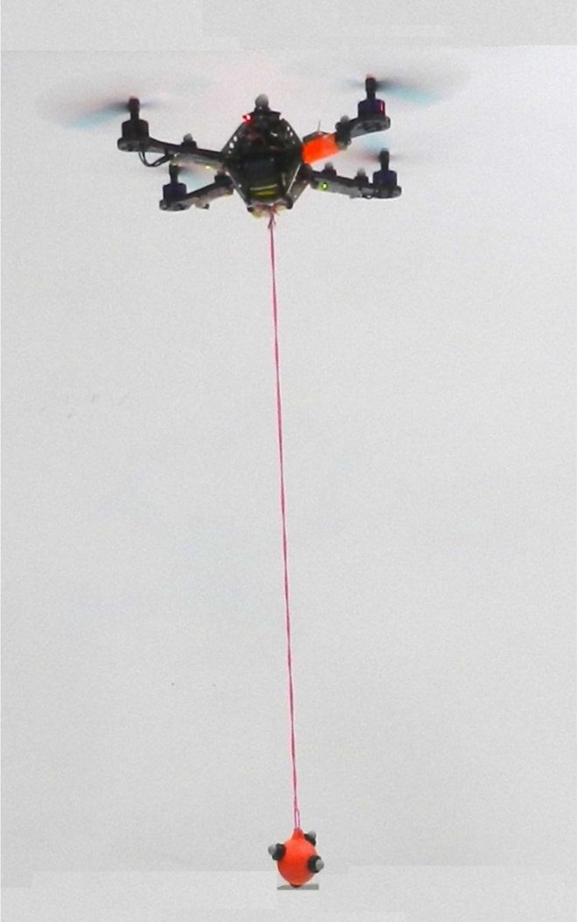
- Aerial transportation applications
  - Humans unable or unwilling
  - Search and rescue missions
  - Supply delivery
  - Patient transportation
  - Wildfires

Photo by UN Photo/Logan Abassi

[1]http://www.statisticbrain.com/wild-fire-statistics/

Problem Formulation

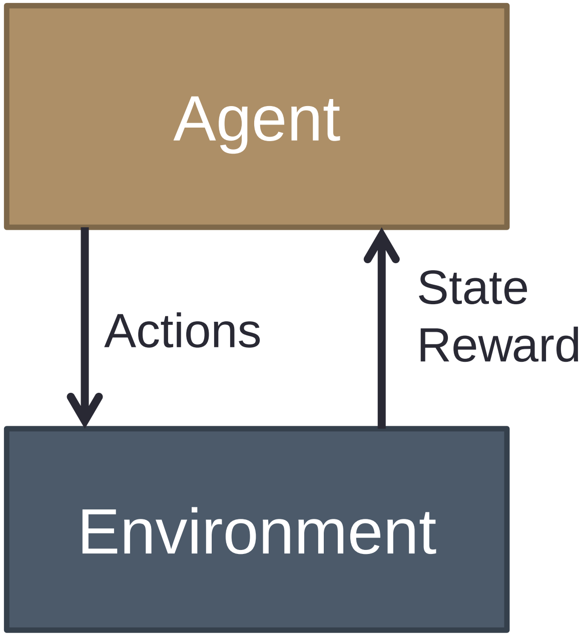
- Holonomic cargo-bearing UAV
  - Bring the suspended load to the destination
  - Minimal residual load oscillations at arrival
- Challenges
  - Non-linear, unknown dynamics
  - Hardware safety
- Dynamical systems balancing constraints problem
  - Quality vs. quantity
  - Anti-lock brakes, traction control



Expert Attempt

Reinforcement Learning (RL)

- Markov decision process MDP (S,A,P,R)
- Value $V:S\text{->}R$
  - Cost to go, potential for accumulated reward
- Induces policy $\pi:S\text{->}A$
  - Action sequence that maximizes the value
  - $\pi(s) = argmax_a V(s')$
  - s' resulting state when a is applied to state s



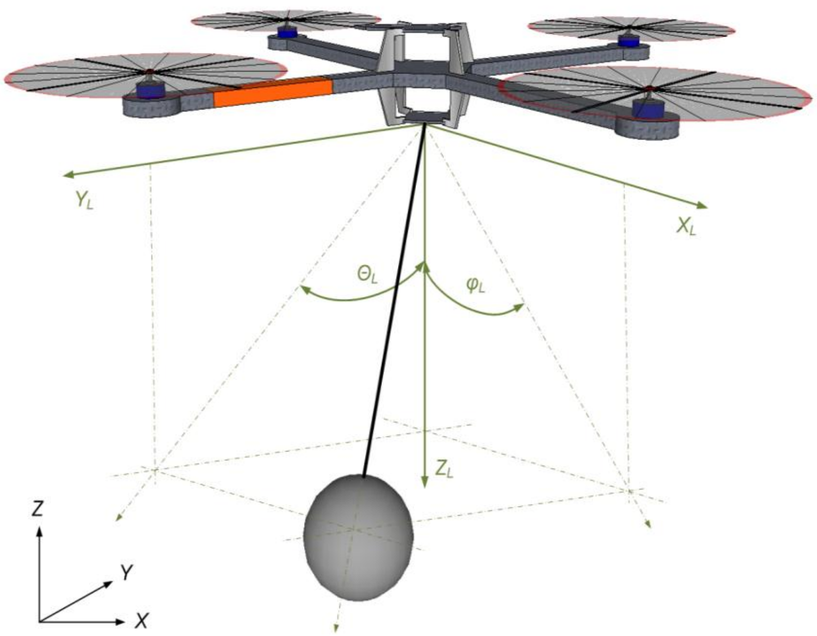
Approximate Value Iteration (AVI)

Ernst et al. 2005

- Offline
- Continuous state, discrete action space
- Iteratively finds value function approximation as
  - $V(s) = \psi^T F(s)$
- Algorithm
  - Sample (s_i, r_i)
  - $v_i = r_i + \gamma * max_{a}\ \psi_n^T F(s_i')$
  - $\psi_{n+1} = argmin_{\psi}\ (v_i - \psi^T F(s_i))^2$

AVI Implementation

- MDP Setup
  - s: vehicle and load position and velocities
  - Discretized acceleration vector
  - Reward structure
  - Generic holonomic vehicle with suspended load
- Goal state in equilibrium
- Learning sampling domain
- Feature vector F(s), squared:
  - Distance from the goal
  - Vehicle's velocity magnitude
  - Load displacement
  - Load’s velocity magnitude



AVI Learning Convergence

- Randomized algorithm
- No convergence guarantees
- Monte Carlo selection
  - Repeat learning over several trials
  - Select best policy
- Can we be sure the policy takes the UAV to the goal?

Trajectory Generation

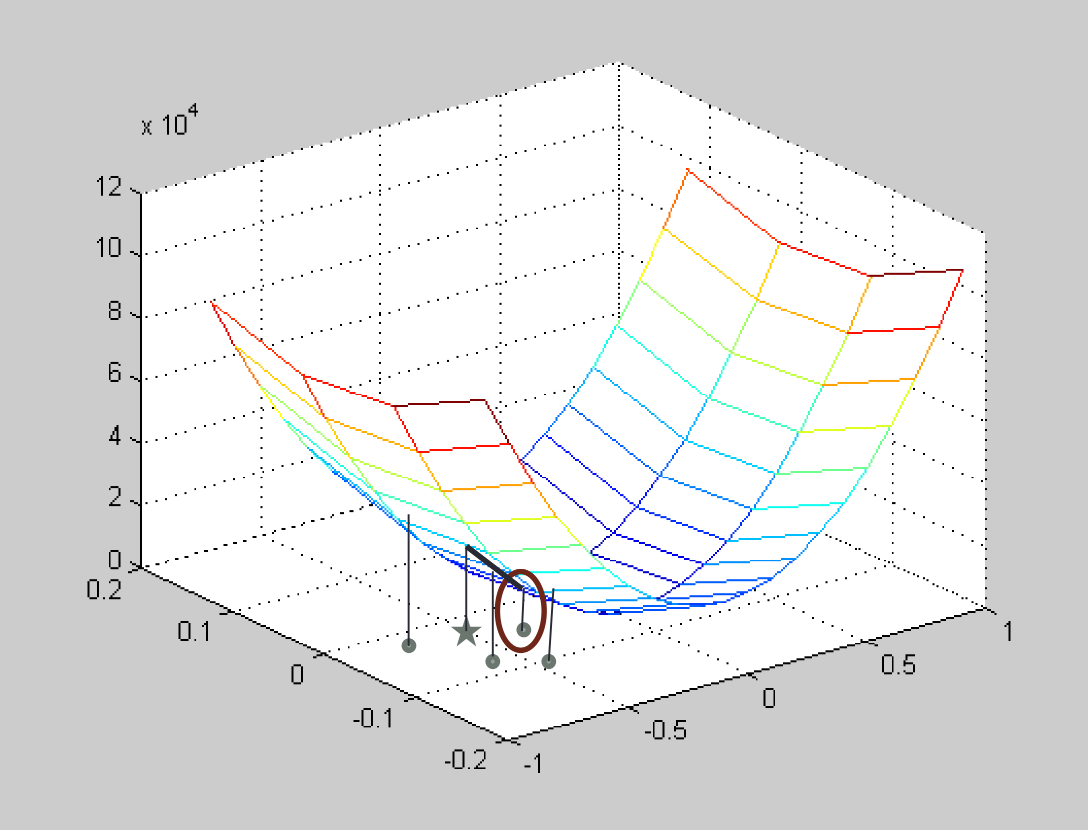
- Lyapunov stability analysis
  - Given start state and control function where will the system end up?
- Control function
  - Policy
  - Transitional probabilities / generative model
  - Action space

Lyapunov Stability Analysis

- $W(s) = -V(s) = -\psi^T F(s)$ is Lyapunov function if:
  - $\psi^i < 0$
  - $\Delta W(s) < 0$
- System is asymptotically stable
  - If action space discretization allows transitions to a higher-valued states for all states
- State space
  - Learning vs. problem domain
  - Lyapunov criteria holds on the problem domain
  - *Learn in small area, and the policy viable starting at arbitrary position (asymptotically stable)*

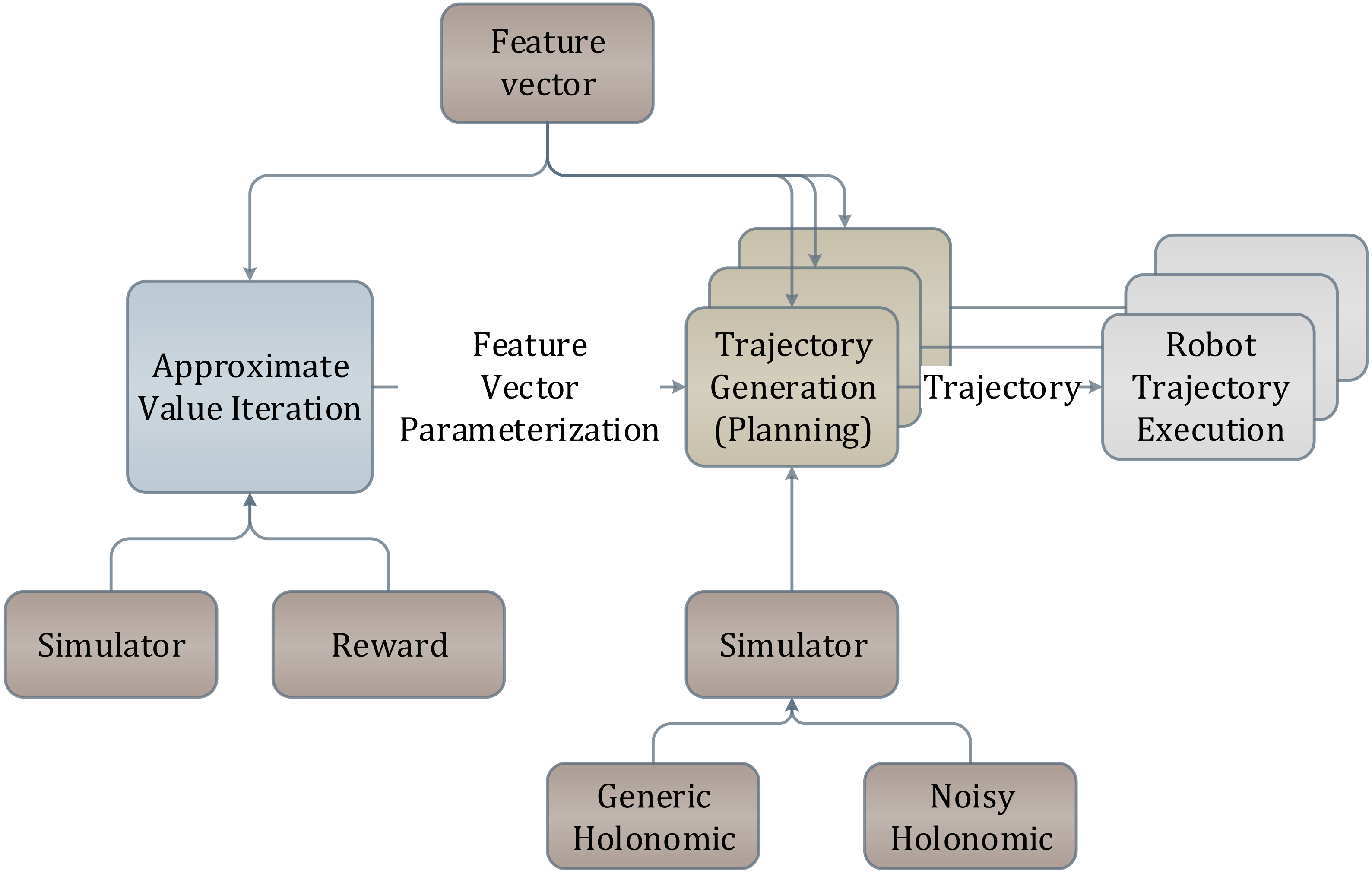
Can we change generative model and action space between learning and planning?

- Generative model and action space
  - Determine state space connectivity
  - Discretize value / Lyapunov function
  - Determine the difference condition of the Lyapunov function



Yes! - as long as we preserve ability to transition to a higher valued / lower cost state

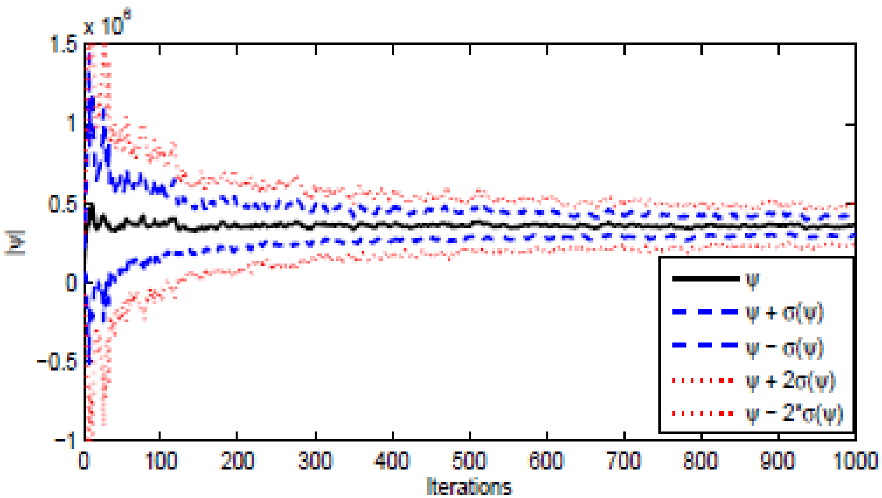
Approach



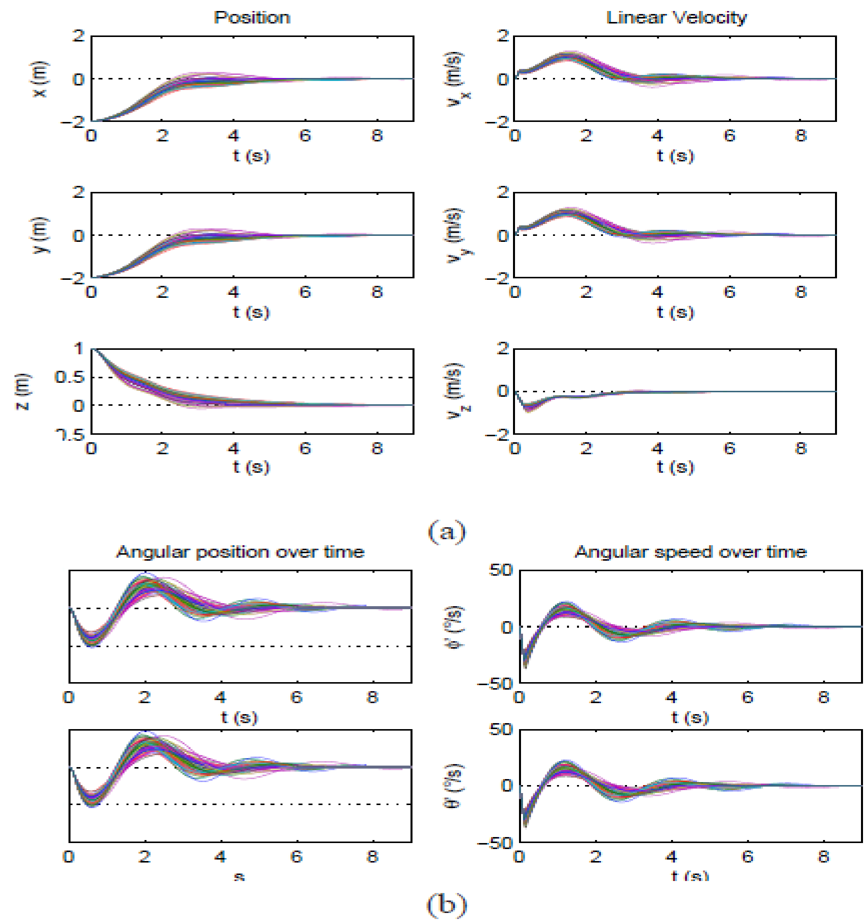
Feature vector
Approximate Value Iteration
Feature Vector Parameterization
Trajectory Generation (Planning)
Trajectory
Robot Trajectory Execution
Simulator
Reward
Simulator
Generic Holonomic
Noisy Holonomic

Learning convergence

- 1000 iterations
- 100 trials
- Fixed start point



Parameter vector norm over 1000 iterations



Trajectories over100 trials from a fixed start point

Learning converges

Policies generate consistent trajectories

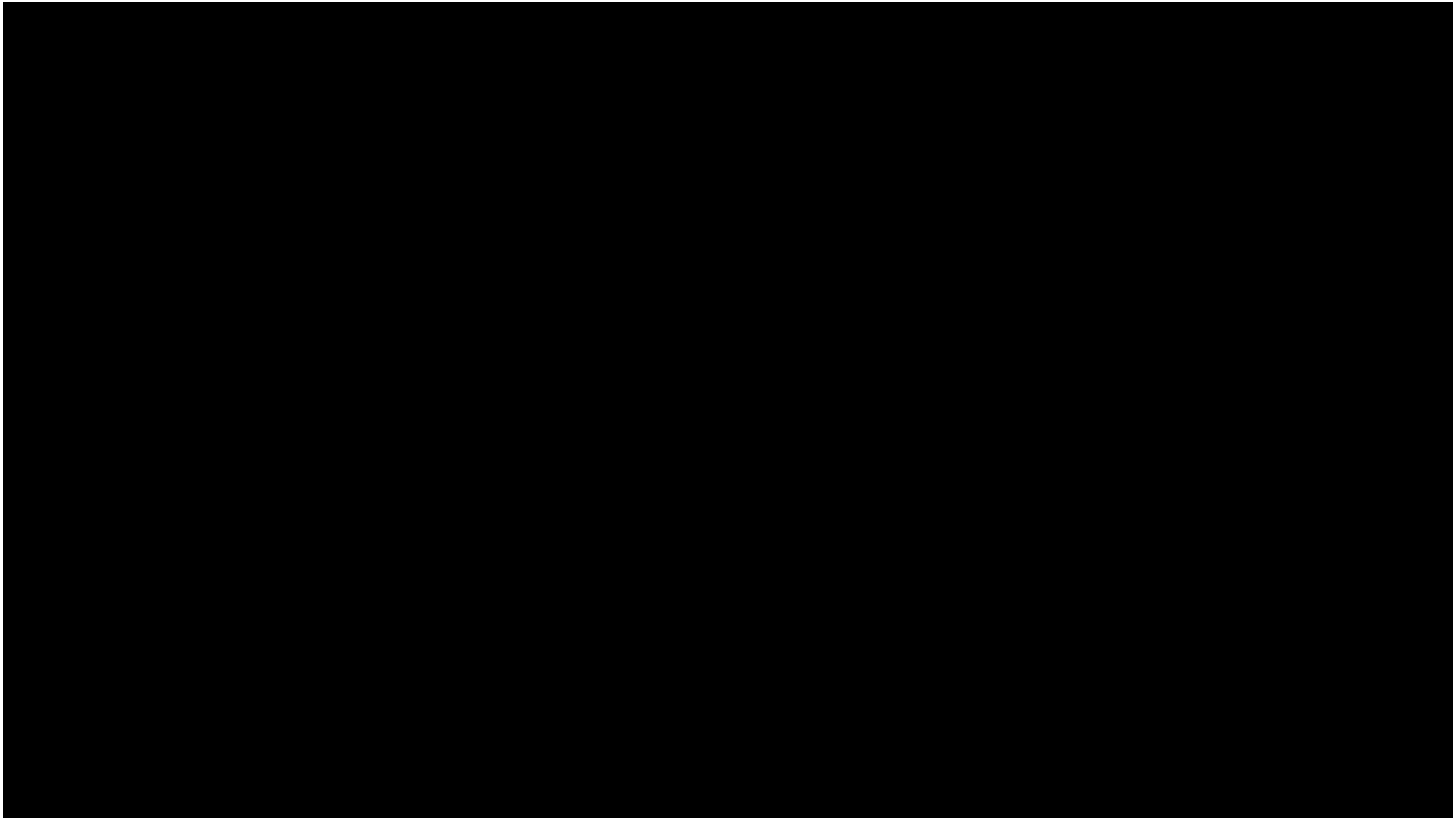
Variable Start Trajectory Generation

- Trained with coarse-grain 3D action set
- Planned with fine-grain 3D action set

State		Goal reached	t (s)		$\parallel p \parallel (m)$		$\parallel \eta \parallel (^\circ)$		$max \parallel \eta \parallel (^\circ)$	
Location	Simulator	(%)	μ	σ	μ	σ	μ	σ	μ	σ
(-2,-2,1)	Generic Holonomic	100	6.13	0.82	0.03	0.01	0.54	0.28	12.19	1.16
	Noisy Holonomic	100	6.39	0.98	0.04	0.01	0.55	0.30	12.66	1.89
(-20,-20,15)	Generic Holonomic	99	10.94	1.15	0.04	0.01	0.49	0.33	46.28	3.90
	Noisy Holonomic	89	12.04	1.91	0.08	0.22	0.47	0.45	44.39	7.22
((4,5),(4,5),(4,5))	Generic Holonomic	100	7.89	0.87	0.04	0.01	0.36	0.31	26.51	2.84
	Noisy Holonomic	100	7.96	1.11	0.04	0.01	0.44	0.29	27.70	3.94
((-1,1),(-1,1),(-1,1))	Generic Holonomic	100	4.55	0.89	0.04	0.01	0.33	0.30	3.36	1.39
	Noisy Holonomic	100	4.55	1.03	0.04	0.01	0.38	0.29	3.46	1.52

Trajectory characteristics for various start positions: duration, ending distance from the goal, ending load swing, and maximum swing encountered.

Trajectory end characteristics are consistent regardless of the start state.



Questions

- Thank you!

- Website:
  - https://www.cs.unm.edu/amprg/Research/Quadrotor/
  - Code, movie, etc…

- Acknowledgements
  - Thanks to Peter Ruymgaart for piloting quadrotor

This work was in part supported by NSF grant ECCS #1027775 and the Department of Energy URPR Grant #DE-FG52-04NA25590. Tapia is supported in part by National Institutes of Health (NIH) Grant P20RR018754 to the Center for Evolutionary and Theoretical Immunology. Palunko is supported by the European Community's Seventh Framework Programme under Grant No. 285939 (ACROSS).

THE UNIVERSITY of NEW MEXICO

Adaptive Motion Planning Research Group

ACROSS
Centre of Research Excellence for Advanced Cooperative Systems